

背景

想到这个题目是因为 @lijiefei 某天跟我说他有师弟面淘宝时被问到“点击率预估的目标到底是什么”，笨狗当时胡乱扯了一通，发现要把这个似乎已经是真理的事情掰清楚还没那么容易，于是有此念想写文一篇详细分析下原因

我和 jiefei 认识是在百度做搜索广告的时候，那就从搜索广告开始说为什么要预估点击率，以及预估点击率的目标。先申明一些名词和假定：

- 1) 每个广告 (Ad) 有一个出价 (Bid), 并有其某情形下实际的点击率 (Click-Through-Rate, CTR)
- 2) 广告按点击收费 (Charge per Click, CPC), 下面我们会分别讨论一价计费 (First-Price, FP, 即广告出价多少则一次点击计费多少) 和二价计费 (Second-Price, SP, 即广告按下一位出价来支付点击价格, 更普遍的是 GSP)
- 3) 千次展现收费 (Cost Per Mille, CPM, 或 RPM, R for Revenue), 即对点击付费广告其展示一千次情况下的收入 (一价计费下等价于 $1000CTR \times Bid$), 或是展示广告的千次展现固定价格
- 4) 预估点击率 (predict CTR, pCTR) 是指对某个广告将要在某个情形下展现前, 系统预估其可能的点击概率

目标分类

搜索广告跟自然结果一个很大的区别就是自然结果只要有一点相关就应该放到所有结果里去, 至于先后位置那个再说, 而广告, 是有个相关性的准入门槛的, 不相关的广告出价再高, 丢出来还是会被骂死. 那怎么判断相关? 用户会用鼠标点击来对结果投票, 相关的广告会被点击, 不相关的广告不会被点击, 那很自然就能得出“点击率和相关性正相关”这个结论 (至于描述里写“二十五岁以下免进”但实际是钢材广告的这种诱骗行为后面再说怎么处理). 那对于这种相关性准入的场景, 预估点击率就是预估广告是否相关, 最朴素情况下这是个二分类问题, 那不管预估成怎样, 只要有一种分割方法能分开是否相关就行了. 此时预估点击率的目标是能对广告按相关与否分类 (或说按相关性排序并给出一个截断值). 评估分类问题好坏, 一般都是看准确和召回两个指标, 用人工打分的记录来做回归验证就行

目标排序

判断相关与否只是点击率预估对广告的一个小辅助, 我们来看看广告的目标是什么? 没错, 是赚钱. (我曾经在其他场合说过广告的目标是维持用户体验下持续赚钱, 不过跟赚钱这一简化目标这不冲突, 前面相关性上已经保证了维持用户体验, 那只要能让广告主还有的赚, 就能持续赚钱) 我们再把问题简化下, 如果广告都是一样的固定价格, 且就以这个价格按点击计费, 那在 PV 一定且预算充分的情况下, 更高的点击率则意味着更赚钱. 这样目标可以等价于怎么挑出更赚钱的广告, 就是那些点击率最高的广告, 我们只要能弄明白广告实际点击率的高低关系就能取得收益最大化, 预估点击率在这时候又是个排序问题, 我们只要弄对广告之间的序关系, 就可以收益最大. 评估排序问题的好坏, 一个经典方法是对 pCTR 的 ROC 曲线算 AUC (曲线下面积), 实际上我见过的做法也都是通过评估 AUC 的高低来判断点击率预估模型的好坏

目标带权排序

上一段里对广告这个业务做了很多简化,比如大家价格都是一样的,如果我们考虑价格不一样的情况,那预期收益就会变成(价格Bid*点击率CTR),这个值很多地方也叫 CPM 或 RPM. 如果是对 CPM 排序,那就需要我们预估的点击率在维持序关系正确的前提下,还要保证相互之间的缩放比是一样的. 比如有广告 A, B, C, 实际点击率是 5%, 3%, 1%, 那在价格一致的情况下,我预估成 5-3-1 还是 5-4-3 是没关系的,但在价格不一样的情况下,比如 1, 1.5, 3, 这时候 5-4-3 的预估点击率值会让他们的预估排名和实际排名刚好颠倒过来,不过预估 5-3-1 或 10-6-2 (放大一倍) 倒没关系. 为了评估这个结果,可以在描 ROC 曲线时把价格乘上去,那最后还是判断排序问题的好坏,加了价格的 AUC 我们可以叫 wAUC (weighted-AUC), 这个离线评估和在线效果依然可以对等

目标准确

从准确召回到 AUC 再到 wAUC, 看起来对已有问题可以完美解决了? 不过广告计费显然不是 FP 这么简单, 在 Google 的带领下几乎所有的搜索引擎都使用了 GSP (Generalized Second Price) 来对广告点击进行计费, 这里再简单解释下最简版 GSP 是怎么回事:

- 1) 所有广告按 CPM 排序 (即 $CTR \times Bid$)
- 2) 排最后的广告收底价 (Reserved Price, RP)
- 3) 其他广告按他的下一位 CPM_{i+1} 除以他的 CTR_i 并加一个偏移量来计费, 并保证比底价高, 即 $Price_i = \max(CPM_{i+1}/CTR_i + \delta_{price}, RP)$

至于 GSP 的细节和为什么这么做能保证收入和体验的平衡等可以详见相关论文, 我们只讨论在 GSP 模式下, 点击率预估的作用和关键点. 根据介绍 GSP 时最后那个公式, 如果把 CPM_{i+1} 拆成 $CTR_{i+1} \times Bid_{i+1}$, 看起来只要保证同比缩放还是会没问题? 但是, 凡事怕但是, 在搜索广告里, 不同的展现位置对点击率还有影响, 比如广告 A, B 在第一位点击率是 5%, 3%, 而在第二位是 3%, 2%, 那只是同比缩放就很难保证最终比较是一致的问题了, 所以最好还是保证预估值跟实际值尽可能接近的好, 这样才能在预估时获得更实际用时完全一样的场景. 评估准确度, 我们有 MAE 和 MSE 等一堆指标, 也是现成的工作的比较好的东西

扩展和吐槽

有行家可能会吐槽说我刚那个不同广告在不同位置的衰减不一致这个说法, 跟公开论文说的不一样, Yahoo 的 paper 里说不同广告在同位置的衰减是一样的. 我只能说, 骚年, 你太天真了... 衰减因子怎么可能只是 $f(pos)$ 这样一个简单函数, 从实际情况来看, 衰减函数和广告是有关的, 但我们又不能对每个广告都去估一个 $f(pos, ad)$, 好在, 我们发现可以把不同的广告做聚类后得到一个 $f(pos, type)$ 的函数簇, 事实上, 最后的衰减函数不仅仅有 pos 和 $type$ 两个因子, 而且里面的因子可以极度简化, 最后的衰减用简单函数就能很好拟合, 我说的够多了, 再说估计要被前东家找麻烦, 你们来感受一下就好

前面也提到介绍的 GSP 是最简版的, 那如果是正常版会有什么不一样? 那就是排序时用的是 $Quality \times Bid$, 这个 Quality 百度叫质量度, 也就是广大广告主望眼欲穿的那几颗星. Quality 和 CTR 有什么不一样? 最简情况下这两个当然是一样的, 但是我们可能还要考虑广告主的信誉度, 目标网页的质量等因素 (比如前面提到过的那种描述欺诈或诱导的广告在这个 Quality 因子中被调整掉), 最后这个 Quality 就会是包含了 CTR 的一个多因子复合值, 那要准确估计这个复合值, 当然也要求其中的每个因子的值尽可能准确. 这里在炒冷饭说准确度, 以及 MAE/MSE 的作用

实际上据我所知各搜索广告平台用的是比正常版的 GSP 还加强或改造过的版本, 里面的因子, 公式, 逻辑更复杂. 在这种情况下还是需要继续强调 CTR 预估的准, 才能做更精确的预估, 从而带来更大的收益

广告之外

呼~ 说完了搜索广告, 我们再简单说下内容广告. 搜索广告几乎都是点击付费, 而内容广告同时存在按点击付费和按展现付费, 那怎么比较一个按点击付费的广告和一个按展现付费的广告哪个预期收益更高? 同样是 CPM, 按展现付费的广告给的是确定值, 而按点击付费的是一个预估值, 通过 $Bid * pCTR$ 得到, 如果 pCTR 不准, 就会导致点击付费广告的预期收益计算不准, 则不一定受益最大. 继续强调预估的准的好处

说完广告我们就能说说其他的, 比如搜索, 比如推荐, 这几个的优化目标如果是带来的量, 因为总体 PV 我们没法人工干预, 且每个点击是等价的, 那最后的优化就变成了点击率, 预估的序关系越接近真实, 可获得的收益更高. 如果不同的点击价值不一样, 那就可以把这个点击换做价格代入广告模型, 因为没有二价计费那么讨厌的变换, 所以就按一价计费来考虑, 保证序正确且等比缩放就能保证收益最大. 如果再激进一点, 评估收益时还加入更复杂的因子而不仅仅是价格这个独立因素, 那自然就要求点击率预估准确, 从而保证做决策时和实际情况一致, 继而保证最终的优化目标最大化

总结

- 1) 点击率预估是为产品的最终目标服务的, 最终目标可以是广告的收入, 广告的相关性, 推荐的接受率等, 看具体场景
- 2) 点击率预估的直接目标根据需求场景不同, 分别是保证预估值和实际值分类正确, 预估序和实际序正确, 预估值和实际值是等比缩放的, 预估值等于实际值
- 3) 要保证离线评估点击率预估的效果, 分别可用分类的准确率和召回率, 排序的 AUC, 带权排序的 wAUC, 相似度 MAE/MSE 来评估。

本文来自笨狗又一窝, [作者snoopy](#).